

FIGURE 1

GGACTAATCTGTGGGAGCAGTTATTCCAGTATCACCCAGGGTGAGCCACACCAGGACTGT
GTTGAAGGGTGTCCCCCTAAATGTAATACCTCCTCATCTTTCTTACACAGTG
TCTGAGAACATTACATTATAGATAAGTAGTACATGGTGGATAACTTCTACTTTAGGAGGA
CTACTCTCTGACAGTCCTAGACTGGTCTTACACTAACAGACACCATGAAGGAGTATGTG
CTCCTATTATTCTGGCTTGTGCTGCCAACCCCTTCTTAGCCCTCACACATCGCACT
GAAGAATATGATGCTGAAGGATAATGGAAGACACAGATGATGATGATGATGATGATG
ATGATGATGATGAGGACAACCTCTTTCCAACAAGAGAGCCAAGAAGCCATTTC
TTTGATCTGTTCCAATGTGCCATTGGATGTCAGTGCTATTCACTGAGTTGTACATTGCTC
AGATTTAGGTTGACCTCAGCCCCAACCAACATTCACTTGATACTCGAATGCTTGATCTC
AAAACAATAAAATTAGGAAATCAAAGAAAATGATTTAAAGGACTCACTTCACTTATGGT
CTGATCCTGAACAACAACAGCTAACGAAGATTCAACCAAAAGCCTTCTAACCAAAAGAA
GTTGCGAAGGCTGTATCTGCCCACAATCAACTAACGTGAAATACCACTTAATCTCCC
CATTAGCAGAACTCAGAATTCACTGAAAATAAGTTAAGAAAATACAAAAGGACACATTCAA
GGAATGAATGCTTACACGTTGGAAATGAGTGCAAACCCCTTCTGATAATAATGGGATAGA
GCCAGGGGCATTGAAAGGGGTGACGGTGTCCATATCAGAATTGCAAGCAGAAACTGACCT
CAGTTCTAAAGGCTTACCAACTTATTGGAGCTTCACTTAGATTATAATAAAATTCA
ACAGTGGAACTTGAGGATTAAACGATAACAAAGAACTACAAAGGCTGGCCTAGGAAACAA
CAAATCACAGATATCGAAAATGGGAGTCTGCTAACATACCACGTGTGAGAGAAATACATT
TGGAAAACAATAAACTAAAAAAATCCCTCAGGATTACAGAGTTGAAATACCTCCAGATA
ATCTTCCTTCATTCTAATTCAATTGCAAGAGTGGGAGTAAATGACTTCTGTCCAACAGTGCC
AAAGATGAAGAAATCTTATACAGTGCAATAAGTTATTCAACAAACCGGTGAAATACTGGG
AAATGCAACCTGCAACATTGCTGTGTTGAGCAGAATGAGTGTTAGCTTGGAAACTTT
GGAATGTAATAATTAGTAATTGGTAATGTCCATTAAATATAAGATTCAAAATCCCTACATT
TGGAAACTTGAACTCTATTAAATAATGGTAGTATTATATAACAGCAAATATCTATTCTCA
AGTGGTAAGTCCACTGACTTATTGACAAGAAATTCAACGGAAATTGCCAAACTATT
GATACATAAGGGGTGAGAGAAACAGCATCTATTGCACTTCTTGTGCTACAAATGAT
CTTACATAATCTCATGCTTGACCATTCTTCTTCTAACAAAAAGTAAGATATTGGTA
TTAACACTTGTATCAAGCACATTAAAAAGAACTGTACTGTAATGGAATGCTTGACT
TAGCAAAATTGCTCTTCATTGCTGTTAGAAAAACAGAATTAAACAAAGACAGTAATGT
GAAGAGTGCATTACACTATTCTTATTCTTAGTAACTGGTAGTACTGTAATATTAAAT
CATCTTAAAGTATGATTGATATAATCTTATTGAAATTACCTTATCATGCTTAGAGCCC
CTTATGTTAAAACAATTCTTAAATAAGCCTCAGTAATGTTCTTACCAACTTGA
TAAATGCTACTCATAGAGCTGGTTGGGCTATAGCATATGCTTTTTTTAATTATT
ACCTGATTAAAAACTCTGTAAAAACGTGTAGTGTTCATAAAATCTGTAACCGCATT
AATGATCCGCTATTATAAGCTTTAATAGCATGAAATTGTTAGGCTATATAACATTGCCAC
TTCAACTCTAAGGAATATTTGAGATATCCCTTGGAAAGACCTGCTTGGAAAGAGCCTGGA
CACTAACAAATTCTACACCAATTGTCCTCAAATACGTATGGACTGGATAACTCTGAGAAA
CACATCTAGTATAACTGAATAAGCAGAGCATCAAATTAAACAGACAGAACCGAAAGCTCTA
TATAATGCTCAGAGTTCTTATGTATTCTTATTGGCATTCAACATATGTAACATCAGAAA
ACAGGAAATTTCATTAAAAATTGGTTGAAAT

FIGURE 2

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34392
<subunit 1 of 1, 379 aa, 1 stop
<MW: 43302, pI: 7.30, NX(S/T): 1
MKEYVLLLFLALCSAKPFFSPSHIALKNMMLKDMEDDDDDDDDDDDDEDNSLFPTREPR
SHFFPFDLFPMCPFGCQCYSRVVHCSDLGLTSVPTNIPFDTRMLDLQNNKIKEIKENDFKGL
TSLYGLILNNNKLTKIHPKAFLTTKKLRRLYLSHNQLSEIPLNLPKSLAELRIHENVKKIQ
KDTFKGMNALHVLEMSANPLDNNGIEPGAFEGVTVFHIRIAEAKLTSVPKGLPPTLLELHLD
YNKISTVELEDFKRYKELQRLGLGNMKITDIENGLANIPRVREIHLENNKLKKIPSGLPEL
KYLQIIFLHSNSIARGVNDFCPTVPKMKKSLYSAISLFNNPVKYWEMQPATFRCVLSRMSV
QLGNFGM
```

Signal sequence.

amino acids 1-15

N-glycosylation site.

amino acids 281-285

N-myristoylation sites.

amino acids 129-135, 210-216, 214-220, 237-243, 270-276, 282-288

Leucine zipper pattern.

amino acids 154-176

FIGURE 3

CGGACGCGTGGCGGACCGTGGGCCCGSGCACCGCCCCGGCCCTCCGCCCTCCGCACTCGCGCTCC
CTCCCTCCGCCGCTCCCGGCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC
GGCCCCGCCGGCCCCCGCTGCTGCTCCTCGGGCTGCTGCTCGCTCCGGCCGGCCCCCGGGCGCCGGCCAGA
GCCCCCCGTGCTGCCATCCGTTCTGAGAAGGAGCCGCTGCCGTTGGAGCGGCAGGCTGCACCTTCGGCG
GAAGGTCTATGCCCTGGACGAGACGTGGCACCCGGACTAGGGCAGGCCATTGGGGTGATGCGCTGCGTGTG
CGCCTGCGAGGCGCCTCAGTGGGTGCGCGTACCAAGGGCCCTGGCAGGGTCAGCTGCAAGAACATCAAACAGA
GTGCCCAACCCCGGCCGTGGGAGCCGAGCTGCCGGACACTGCTGCCAGACCTGCCAGTTAGGGCAGCTGACAGGCG
CAGTTCGGAGCGCAGCCGAGCGGCCCTGCTCCTCGAGTATCCGCGGACCCGGAGCATGCCAGTTAGGGCAGCTG
CGGGGAGCCAGGCGCTGAGGAGCGGGCCCTGGTGACGCCACACGGACTTCGTGGCGCTGCTGACAGGGCGAG
GTCGCAAGCGGTGGCACAGAGCCGAGTCTCGCTGCGCTTAGCCTCCGCTCTATCTCCTACAGGCGGCT
GGACCGCCCTACAGGATCCGCTTCAGACTCCAATGCCAGTGTCTGTTGAGCACCCCTGCAGCCCCCACCA
AGATGGCCTGGTCTGTGGGTGTGGCGGGAGTGCCTCGGTTGTCTCTGCGGCTCCTAGGGCAGAACAGCTGCA
TGTGGCACTTGTGACACTCACTCACCCCTCAGGGGAGGTCTGGGGCCCTCTCATCCGGCACCGGGCCCTGGCTGC
AGAGACCTTCAGTGCCATCCTGACTCTAGAAGGCCCCCACAGCAGGGCGTAGGGGGCATCACCTGCTCACTCT
CAGTGACACAGAGGACTCCTGCTTTGCTGCTCTCCGAGGGCTGCTGGAACCCAGGAGTGGGGACTAAC
CCAGGTTCCCTTGAGGCTCCAGATTCTACACCAGGGCAGCTACTGCGAGAACCTCAGGCCATGTCTCAGCCA
GGAACCAGGCTTGCTGAGGTGCTGCCAACCTGACAGTCCAGGAGATGGACTGGCTGGCTGGGAGCTGCA
GATGGCCCTGGAGTGGCAGGCAGGCCAGGGCTGCGCATCAGTGGACACATTGCTGCCAGGAAGAGCTGCGACGT
CCTGCAAAGTGTCCCTTGAGGCTGATGCCCTGATCCCAGTCCAGACGGGTGCTGCCGCTCAGCCAGCCTCAC
GCTGCTAGGAAATGGCTCCCTGATCTATCAGGTGCAAGTGGTAGGGACAAGCAGTGGAGTGGCTGGCATGACACT
GGAGACCAAGCCTCAGCGGAGGGATCAGCGCACTGCTCTGTGCCACATGGCTGACTCCAGCCAGGAGGACACAC
GGCCGTGGGTATCTGCCCTGGCTGGTGCCGAGGGCTCATATGCTGCTGCAAGTGGACTCTTCTGAAACGT
GGGCACCAAGGACTTCCCAGACGGAGAGCTCGGGGCACTGGCTGCCCTACTGTGGCATAGGCCCG
CCATGACACGCTGCCGTGCCCTAGCAGGCCCTGGTGTACCCCTGTGAAGAGCCAAGCAGCAGGGCACGC
CTGGCTTCCCTGGATAACCCACTGTCACCTGCACTATGAAGTGTGCTGGCTGGCTGGCTGGCTCAGAACAAAG
CACTGTCACTGCCAACCTCCTGGCTGGTGAAACGCCAGGGCCTCGGCCGCTGTAAGGGATTCTATGGCTC
AGAGGCCAGGGTGTGGTAAGGACCTGGAGCCGAAGTGTGCGGCACCTGGCAAAAGGCATGGCTCCCTGAT
GATCACCACCAAGGGTAGCCCCAGAGGGAGCTCGAGGGCAGGTGACATAGCCAACCAATGTGAGGTTGGCGG
ACTGCGCCTGGAGGCCGGCGAGGGGTGCGGGCGCTGGGCTCCGATACAGCCTCTGCTGCCGCG
TGTGGTGCCTGGTCTCCGGCCCTAGGCCGCCAAACCTGGTGTCTGGCGGCCAGACCCCAACACATG
CTTCTCGAGGGCAGCAGGCCAACGGGGCTGCTGGCGCCAACCTACGACCCGCTCTGCTCACTCTGCA
CTGCCAGAGACGAACGGTGAATGTGACCCGGTGGTGTGCCACGCCAGCTGCCAACACCGGTGCAAGGCTCC
CGACCAAGTGTGCCCTGTTGCCCTGAGAAAACAAGATGTCAAGAGACTTGGCCAGGGCTGCCAAGGAGCCGGACCC
AGGAGAGGGCTGCTATTGTGACCGGAGCTGGCGGGCAGCGGGTACCGGGTGGCACCCGTTGTGCCCC
CTTGGCTTAATTAAGTGTGCTGTCACCTGCAAGGGGGCACTGGAGAGGTGCACTGTGAGAACGGTGCAGTG
TCCCCGGCTGCCCTGTGCCAGCCTGTGCGTGTCAACCCACCGACTGCTGCAAACAGTGTCCAGTGGGTG
GGCCCAACCCAGCTGGGGACCCATGCAAGGCTGATGGGCCGGCTGCCGTTTGTGCTGGCAGTGGTCCC
AGAGAGTCAGAGCTGGCACCCCTCAGTGGCCCTTTGGAGAGATGAGCTGTATCACCTGCAAGATGTGGGGCAGG
GGTGCCTCACTGTGAGGGGATGACTGTTCACTGCCACTGTGCTGTGGCTGGGAAGGAGAGTCGATGCTGTT
CCGCTGCACGCCAACCGCGGCCAGAGACAGAACACTGATCCAGAGCTGGAGAAAGAACCGGAAGGCTCTA
GGGAGCAGCCAGAGGGCAAGTGAACCAAGAGGATGGGCTGAGCTGGGAAGGGGTGGCATCGAGGACCTTCTT
GCATTCTCCTGTGGGAAGCCCAGTGCCTTGTCTCTGCTCTACTCCACCCCAACTACCTCTGGGAA
CCACAGCTCCACAAGGGGGAGAGGCAGCTGGGCCAGACCGAGGTACAGCCACTCCAAGTCTGCCCTGCCACCC
TCGGCTCTGTGCTGGAGCCCCACCCCTTCTGTACATAATGTCAGTGGCTGTGGATTTTAATT
TCTTCACTCAGCACCAAGGGCCCCGACACTCCACTCTGCTGCCCTGAGCTGAGCAGAGTCATTATTGGAGAG
TTTGTATTATTAAAACATTCTTTCAAGTCAAAAAAAAAAAAAAA

FIGURE 4

><subunit 1 of 1, 954 aa, 1 stop

><MW: 101960, pI: 8.21, NX(S/T): 5

MPSLPAPPAPLLLLGSRPARAGAGPEPPVLPIRSEKEPLPVRAAGCTFGGVYALDE
TWHPDLGQPFGVMRVCVLCAEAPQWGRRTRGPGRVSKNIKPECPTPACGQPRQLPGHCCQT
CPQERSSSERQPSGLSFEYPRDPEHRSYSDRGEPGAERARGDHTDFVALLTGPRSQAVER
ARVSLLRSSLRFSISYRRLLDRPTRIRFSDSNGSVLFEPAAAPTQDGLVCGVWRAPRLSLRL
LRAEQLHVALVTLTHPSGEVWGPLICRHRALAAETFSAILTLEGPQQGVGGITLLTLSDTED
SLHFLLLFRGLLEPRSGGLTQVPLRLQILHQGQLLRELQANVSAQEPGFAEVLPNLTQEMD
WLVLGELOMALEWAGRPGRLISGHIAARKSCDVLQSVLCGADALIPVQTAAGSASLTLLGN
GSLIYQVQVVGTSSEVVAMTLETKPQRRDQRTVLCHMAGLQPGGHTAVGICPGLGARGAHML
LQNELFLNVGKDFPDGELRGHVAALPYCGHSARHDTPVPLAGALVLPPVKSQAAGHAWLS
LDTHCHLHYEVLLAGLGGSEQGTVTAHLLGPPGTPGPRRLLKGFYGSEAQGVVKDLEPELLR
HLAKGMASLMITTGSPRGELRGQVHIANQCEVGLRLEAAGAEGVRALGAPDTASAAPPVV
PGLPALAPAKPGGPGPRDPNTcffEGQQRPHGARWAPNYDPLCSLCTCQRRTVICDPVVCP
PPSCPQVQADQCCPVCPEKQDVRLPGLPRS RDGEGCYFDGDRSWRAAGTRWHPVVPPF
GLIKCAVCTCKGGTGEVCEVQCPRLACAQPVRVNPTDCKQCPVGSGAHPQLGDPMQADG
PRGCRCFAGQWFPEQSWSHPSVPPFGEMSCITCRCAGVPHCERDDCSLPLSCGSGKESRCCS
RCTAHRRPPETRTDPELEKEAEGS

Signal sequence.

amino acids 1-23

N-glycosylation sites.

amino acids 217-221, 351-355, 365-369, 434-438

Tyrosine kinase phosphorylation sites.

amino acids 145-153, 778-786

N-myristoylation sites.

amino acids 20-26, 47-53, 50-56, 69-75, 73-79, 232-238, 236-242,
390-396, 422-428, 473-479, 477-483, 483-489, 489-495, 573-579,
576-582, 580-586, 635-641, 670-676, 773-779, 807-813, 871-877,
905-911

Amidation site.

amino acids 87-91

Cell attachment sequence.

amino acids 165-168

Leucine zipper pattern.

amino acids 315-337

FIGURE 5

GGCGGAGCAGCCCTAGCCGCCACCGTCGCTTCGCAGCTCTCGTCGCCACTGCCACCGCCGCCGTCACTGCG
TCCTGGCTCCGGCTCCCGGCCCTCCGGCCGGC**ATG**CAGCCCCGCCGCCAGGCGCCCGGTGCGCAGCTGC
TGCCCGCGCTGGCCCTGCTGCTGCTGCTGGAGCAGGGGGCCCCGAGGCAGCTCCCTGGCCAACCCGGTGCCCG
CCCGGCCCTTGCTCGGCCCGGGCGTGCAGCCCTGCCGAATGGGGGTGTGTCACCTCGCGCCCTG
AGCCGGACCCGAGCACCCGGCCCCGCCGGAGCCTGGCTACAGCTGCACCTGCCCGCCGGATCTCCGGCG
CCAAGTGCAGCTTGTGAGATCCTTGCCAGCAACCCCTGTCAACATGGCAACTGCAGCAGCAGCA
GCAGCAGCGATGGCTACCTCTGCAATTGCAATGAAGGCTATGAAGGTCCAACGTGAACAGGCACTTCCCAGTC
TCCCAGCCACTGGCTGGACCGAATCCATGGCACCCGACAGCTTCAGCCTGTTCTGCTACTCAGGAGCCTGACA
AAATCCTGCCTCGCTCTCAGGCAACGGTGACACTGCCTACCTGGCAGCCAAAACAGGGCAGAAAGTTGAGAAA
TGAAATGGGATCAAGTGGAGGTGATCCCAGATATTGCGCTGTGGGAATGCCAGTTCTAACAGCTCTGGGGTGGCC
GCCTGGTATCCTTGAAAGTGCACAGAACACCTCAGTCAAGATTGGCAAGATGCCACTGCCTCACTGATTTC
TCTGGAAGGTACGGCCACAGGATTCCAACAGTGCCTCCATAGATGGACGAAGTGTGACCCCCCTCAGGCTT
CAGGGGACTGGCTCTGGAGGAGATGTCGCTGGCTGGGAATAACTTATTGGTTTGATGATTCTG
TGACTAAAGTCTATTGTGGCTTGCGCTTAACCTGGTGGTGAAGGTCAAGCACCTGTGTGCCGGGGAGAGTCACG
CAAATGACTTGGAGTGTTCAGGAAAAGGAAAATGACACCAGAAGCCCTCAGAGGCAACTTTTCTGTACCTGT
AGGAGCAGTACGTGGTACTTCTGTGAAGAATACGATGCTTGCAGAGGAAACCTGCCAAAACAACGCGAGCT
GTATTGATGCAAATGAAAAGCAAGATGGGAGCAATTTCACCTGTGTTGCTCTGGTTACTGGAGAGCTTT
GCCAGTCAAAGATTGATTACTGCATCCTAGACCCATGCAAGAAATGGAGCAACATGCAATTCCAGTCTCAGTGGAT
TCACCTGCCAGTGTCCAGAAGGATACTTCGGATCTGCTGTGAAGAAAAGGTGGACCCCTGCGCCTCGTCTCCGT
GCCAGAACACGGCACCTGCTATGGACGGGTACACTTACCTGCAACTGCAGCCCAGGCTTCACAGGGCCGA
CCTGTGCCAGCTTATTGACTTCTGTGCCCTCAGCCCCTGTGCTCATGGCACGTGCCGAGCGTGGCACAGCT
ACAAATGCCCTGTGATCCAGGTACCATGGCTCTACTGTGAGGAGGAATATAATGAGTGCCTCTCCGCTCCAT
GCCTGAATGCCACCTGCAGGGACCTCGTTAATGGCTATGAGTGTGTGCCAGAATACAAAGGAACAC
ACTGTGAATTGTACAAGGATCCCTGCGCTAACGTCAGCTGTGAACGGAGCCACCTGTGACAGCGACGGCCTGA
ATGGCACGTGCATCTGTGACCCGGTTACAGGTGAAGAGTGCAGCATTGACATAATGAATGTGACAGTAACC
CCTGCCACCATGGTGGAGCTGCCCTGGACCAGCCAATGGTTATAACTGCCACTGCCGATGGTTGGGTGGAG
CAAACGTGAGATCCACCTCAAATGGAAGTCCGGGCACATGGCGAGAGCCTACCAACATGCCACGGCACTCCC
TCTACATCATCATTGGAGCCCTCTCGTGGCCTTCATCCTATGCTGATCATCCTGATCGTGGGATTGCGC
TCAGCCGCATTGAATACCAGGGTTCTCCAGGCCAGCCTATGAGGAGTTCTACAACGTGCCGCAGCATCGACAGCG
AGTTCAAGCAATGCCATTGCACTCCATCGGCATGCCAGGTTGGAAAGAAATCCGGCCTGCAATGTATGATGTGA
GCCCATGCCCTATGAAGATTACAGTCTGTGACAAACCCCTGGTCACACTGATTTAAACTAAAGATTG**TAA**
CTTTTTGGATTATTTCAAAAAGATGAGATACTACACTCATTAAATATTTAAAGAAAATAAAAGCTTAA
GAAATTAAATGCTAGTGTCAAGAGTTTCAGTAGAAATTTAAGAACTAATTTCTGCAGCTTTAGTTG
GAAAAATATTTAAAACAAAATTGTGAAACCTATAGACGATTTAATGTACCTCAGCTCTAAACTGT
GTGCTCTACTAGTGTGTGCTCTTCACTGTAGACACTATCACGAGACCCAGATTAACTTCTGTGGTTGTTACA
GAATAAGTCTAATCAAGGAGAAGTTCTGTTGACGTTGAGTGCAGGCTTCTGAGTAGAGTTAGAAAACCAC
GTAACGTAGCATATGATGTATAATAGAGTATACCGTTACTTAAAAGAAGTGTGAAATGTTGTTGTGGAA
AGAAAATAGTTAAATTTACTATTCTAACCCGAATGAAATTAGCCTTGCCTTATTCTGTGCATGGTAAGTAAC
TTATTCTGCACTGTTGTTGAACTTGTGAAACATTCTTCAGTTGTTGTGATTTCTGTAACAGTC
TCGAACTAGGCCCTAAAAACATACGTAACGAAAAGGCCAGCGAGGCAAATTCTGATTGATTGAATCTATATT
TTCTTAAAAGTCAGGGTTCTATATTGTGAGTAAATTAAATTTACATTGAGTTGTTGCTAAGAGGTAG
TAAATGTAAGAGAGTACTGGTCCCTCAGTAGTGTGAGTATTCTCATAGTGCAGCTTATTATCTCCAGGATGTT
TTTGTGGCTGTATTGATTGATATGTGCTTCTGATTCTGCTAATTCCAACCATAATTGAATAATGTGATC
AAGTCA

FIGURE 6

><subunit 1 of 1, 737 aa, 1 stop

><MW: 78475, pI: 5.09, NX(S/T): 11

MQPRRAQAPGAQQLPALLLLLLGAGPRGSSLANPVPAAPLSAPGPCAAQPCRNGGVCTSR
PEPDHQHPAPAGEPGYSCTCPAGISGANCQLVADPCASNPCHGNCSSSSSSDGYLCICN
EGYEGPNCEQALPSLPATGWTESMAPRQLQPVPATQEFDKILPRSQATVTLPTWQPKTGQKV
VEMKWDQVEVIPDIACGNASSNSSAGGRLSFEVPQNTSVKIRQDATASLILLWKVTATGFQ
QCSLIDGRSVTPLQASGGLVLLLEMLALGNNHFIGFVNDSVTKSIVALRTLTVVKVSTCVPG
ESHANDLECSGKGKCTKPSEATFSCTCEEQYVGTCEEYDACQRKPCQNNASCIDANEKQD
GSNFTCVCLPGYTGELCQSKIDYCILDPCRNGATCISSLSGFTQCPEGYFGSACEEKVDPC
ASSPCQNNGTCYVDGVHFTCNCSPGFTGPTCAQLIDFCALSPCAHGTCSRSGTSYKCLCDPG
YHGLYCEEYNECLSAPCLNAATCRDLVNGYECVCLAELYKGTHCELYKDPCANVSCLNGATC
DSDGLNGTCICAPGFTGEEDIDINECDSNPCHGGSCLDQPNGYNCHCPHWVGANCEIHL
QWKSGHMAESLTNMPRHSLYIIGALCVAFILELIIVGICRISRIEYQGSSRPAYEEFYN
CRSIDSEFSNAIASIRHARFGKKSRPAMYDVSPIAYEDYSPDDKPLVTLIKTKDL

Signal sequence.

amino acids 1-28

Transmembrane domain.

amino acids 641-660

N-glycosylation sites.

amino acids 107-111, 204-208, 208-212, 223-227, 286-290, 361-365,
375-379, 442-446, 549-553, 564-568

Glycosaminoglycan attachment site.

amino acids 320-324

Tyrosine kinase phosphorylation sites.

amino acids 490-498, 674-682

N-myristoylation sites.

amino acids 30-36, 56-62, 57-63, 85-91, 106-112, 203-209,
373-379, 449-455, 480-486, 562-568, 565-571

Amidation site.

amino acids 702-706

Aspartic acid and asparagine hydroxylation site.

amino acids 520-532, 596-608

EGF-like domain cysteine pattern signatures.

amino acids 80-92, 121-133, 336-348, 378-390, 416-428, 454-466,
491-503, 529-541, 567-579, 605-617

FIGURE 7

CTCTGGAAGGTACGGCCACAGGATTCCAACAGTGCTCCCTCATAGATGGACGAAAGTGTGA
CCCCCCTTCAGGCTTCAGGGGGACTGGTCCTCCTGGAGGAGATGCTCGCCTGGGAATA
ATCACTTTATTGGTTTGTGAATGATTCTGTGACTAAGTCTATTGTGGCTTGCCTTAAC
CTGGTGGTGAAGGTCA~~G~~CACCTGTGTGCCGGGGAGAGTCACGCAAATGACTTGGAGTGTTC
AGGAAAAGGAAAATGCACCACGAAGCCGT~~C~~AGAGGCAACTTTCTGTACCTGTGAGGAGC
AGTACGTGGGTACTTCTGTGAAGAATACGATGCTGCCAGAGGAAACCTGCCAAACAAAC
GCGAGCTGTATTGATGCAAATGAAAAGCAAGATGGAGCAATT~~T~~CACCTGTGTTGCCTTCC
TGGTTATACTGGAGAGCTTGCCAACCGAACTGAGATTGGAGCGAACGACCTACACCGAACT
GAGATAGGGAG

FIGURE 8

CTCTGGAAGGTACGGCCACAGGATTCCAACAGTGCTCCCTCATAGATGGACGAAAGTGTGA
CCCCCCTTCAGGCTTCAGGGGGACTGGTCCTCCTGGAGGAGATGCTCGCCTGGGAATA
ATCACTTTATTGGTTTGTGAATGATTCTGTGACTAAGTCTATTGTGGCTTGCCTTAAC
CTGGTGGTGAAGGTCAAGCACCGTGTGCCGGGGAGAGTCACGCAAATGACTTGGAGTGTTC
AGGAAAAGGAAAATGCACCACGAAGCCGTCAAGAGGCAACTTTCTGTACCTGTGAGGAGC
AGTACGTGGGTACTTCTGTGAAGAATACGATGCTGCCAGAGGAAACCTGCCAAAACAAC
GCGAGCTGTATTGATGCCAATGAAAAGCAAGATGGAGCAATTACCTGTGTTGCCTTCC
TGGTTATACTGGAGAGCTTGCCAACCGAACTGAGATTGGAGCGAACGACCTACACCGAACT
GAGATAGGGGAG

FIGURE 9

GCTGAGTCTGCTGCTCCTGCTGCTGCTCCAGCCTGTAACCTGTGCCTACACCACGCCAG
GCCCCCCCAGAGCCCTCACCAACGCTGGCGCCCCAGAGCCCACACC**ATG**CCGGGCACCTAC
GCTCCCTCGACCACACTCAGTAGTCCCAGCACCCAGGGCCTGCAAGAGCAGGCACGGCCCT
GATGCGGGACTTCCCGCTCGTGGACGCCACAACGACCTGCCCCTGGCCTAAGGCAGGTTT
ACCAGAAAGGGCTACAGGATGTTAACCTGCGCAATTTCAGCTACGCCAGACCAGCCTGGAC
AGGCTTAGAGATGGCCTCGTGGCGCCAGTTCTGGTCAGCCTATGTGCCATGCCAGACCCA
GGACCGGGATGCCCTGCGCCTCACCCCTGGAGCAGATTGACCTCATGCCGATGTGTGCCT
CCTATTCTGAGCTGGAGCTTGACCTCGGCTAAAGCTCTGAACGACACTCAGAAATTGGCC
TGCCTCATCGGTGTAGAGGGTGGCCACTCGCTGGACAATAGCCTCTCCATCTTACGTACCTT
CTACATGCTGGAGTGCCTACCTGACGCTACCCACACCTGCAACACACCCCTGGCAGAGA
GCTCCGCTAAGGGCGTCCACTCCTCTACAACAAACATCAGCGGGCTGACTGACTTTGGTGAG
AAGGTGGTGGCAGAAATGAACCGCCTGGCATGATGGTAGACTTATCCCATGTCTCAGATGC
TGTGGCACGGCGGGCCCTGGAAGTGTACAGGCACCTGTGATCTCTCCACTCGCTGCC
GGGGTGTGTGCAACAGTGCTCGGAATGTTCTGATGACATCCTGAGCTCTGAAGAAGAAC
GGTGGCGTCGTGATGGTGTCTTGCCATGGAGTAATACAGTGCAACCCATCAGCCAATGT
GTCCACTGTGGCAGATCACTCGACCACATCAAGGCTGTCATTGGATCCAAGTTCATCGGA
TTGGTGGAGATTATGATGGGCCGCAAATTCCCTCAGGGGCTGGAAGACGTGTCCACATAC
CCGGTCCTGATAGAGGAGTTGCTGAGTCGTGGCTGGAGTGAGGAAGAGCTTCAGGGTGTCC
TCGTGGAAACCTGCTCGGGCTTCAGACAAGTGGAAAAGGTACAGGAAGAAAACAAATGGC
AAAGCCCCTGGAGGACAAGTTCCCGATGAGCAGCTGAGCAGTTCTGCCACTCCGACCTC
TCACGTCTGCGTCAGAGACAGAGTCTGACTTCAGGCCAGGAACACTCACTGAGATTCCCATA
CTGGACAGCCAAGTTACCAAGCCAAGTGGTCAGTCTCAGAGTCCTCCCCCACATGGCCCCAG
TCCTTGCAGTTGTGGCCACCTTCCCAGTCCTTATTCTGGCTCT**TGA**TGACCCAGTTAGTCC
TGCCAGATGTCACTGTAGCAAGCCACAGACACCCCACAAAGTTCCCTGTTGTGCAGGCACA
AATATTTCTGAAATAATGTTTGGACATAG

FIGURE 10

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA35595
<subunit 1 of 1, 433 aa, 1 stop
<MW: 47787, pI: 6.11, NX(S/T): 5
MPGTYAPSTTLSSPSTQGLQEQRALMRDFPLVDGHNDLPLVLRQVYQKGLQDVNLRNFSYG
QTSLDRLRDGLVGAQFWSAYVPCQTQDRDALRLTLEQIDLIRRMCA SYSELELVTSAKALND
TQKLACLIGVEGGHSLDNSLSILRTFYMLGVRYLTHTCNPWAESSAKGVHSFYNNISGL
TDFGEKVVAEMNRLGMMVDLSHVSDAVARALEVSQAPVIFSHSAARGVCNSARNVPDDILQ
LLKKNGGVMVSLSMGVIQCNPSANVSTVADHFDHIKAVIGSKFIGIGG DYDGAGKFPQGLE
DVSTYPVLI EELLSRGWSEEELQGVLRGNLLRVFRQVEKVQEEWKWQS PLEDKFPDEQLSSS
CHSDL SRLRQRQSLTSGQELTEIPIHWTAKLPAKWSVSESSPHMAPVLA VVATFPV LILWL
```

N-glycosylation sites.

amino acids 58-62, 123-127, 182-186, 273-277

N-myristoylation sites.

amino acids 72-78, 133-139, 234-240, 264-270, 334-340, 389-395

Renal dipeptidase active site.

amino acids 134-157

FIGURE 11

AAAACCTATAAATATTCCGGATTATTCAACCGTCCCACCATGGGCGGGATCCGCAGGCC
CGAATTCTAAACCAACATGCCGGGCACCTACGCTCCCTCGACCACACTCAGTAGTCCCAGCA
CCCAGGGCCTGCAAGAGCAGGCACGGGCCTGATGCGGGACTTCCGCTCGTGGACGGCCAC
AACGACCTGCCCTGGTCCTAAGGCAGGTTACCAGAAAGGGCTACAGGATGTTAACCTGCG
CAATTCAGCTACGCCAGACCAGCCTGGACAGGCTTAGAGATGCCCTCGTGGCGCCAGT
TCTGGTCAGCCTATGTGCCATGCCAGACCCAGGGATGCCCTCGCCTCACCTGGAG
CAGATTGACCTCATGCCGCATGTGCGCTCCTATTCTGAGCTGGAGCTTGTGACCTCGC
TAAAGCTCTGAACGACACTCAGAAATTGGCCTGCCTCATCGGTGTAGAGGGTGGCACTCGC
TGGACAATAGCCTCTCCATCTTACGTACCTTACATGCTGGGAGTGCCTACCTGACGCTC
ACCCACACCTGCAACACACCCCTGGCAGAGAGCTCCGCTAAGGGCGTCCACTCCTTACAA
CAACATCAGCGGGCTGACTGACTTTGGTGAGAAGGTGGCAGAAATGAACCGCCTGGCA
TGATGGTAGACTTATCCCATGTCTCAGATGCTGTGGCACGGCGGGCCCTGGAAGTGTACAG
GCACCTGTGATCTCTCCCACTCGGCTGCCCGGGGTGTGCAACAGTGCTCGGAATGTTCC
TGATGACATCCTGCAGCTTCTGAAGAAGAACGGTGGCGTCGTGATGGTGTCTTGCCATGG
GAGTAATAACAGTGCAACCCATCAGCCAATGTGTCACGTGGCAGATCACTCGACCACATC
AAGGCTGTATTGGATCCAAGTTCATCGGATGGTGAGATTATGATGGGCCGCAAATT
CCCTCAGGGCTGGAAGACGTGTCACATACCCGGCCTGATAGAGGAGTTGCTGAGTCGTG
GCTGGAGTGAGGAAGAGCTTCAGGGTGTCTCGTGGAAACCTGCTGCGGGCTTCAGACAA
GTGGAAAAGGTACAGGAAGAAAACAAATGGCAAAGCCCTGGAGGACAAGTTCCGGATGA
GCAGCTGAGCAGTCCCTGCCACTCCGACCTCTCACGTCTCGTCAGAGACAGAGTCTGACTT
CAGGCCAGGAACACTGAGATTCCCACACTGGACAGCCAAGTTACCAGCCAAGTGGTCA
GTCTCAGAGTCCTCCCCCACCCTGACAAAACTCACACATGCCACCGTGCCAGCACCTGA
ACTCCTGGGGGACCGTCAGTCTCCTCTTCCCCCAAAACCCAAGGACACC

FIGURE 12

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA35872
><subunit 1 of 1, 446 aa, 0 stop
><NX(S/T): 5
MPGTYAPSTTLSSPSTQGLQEQRALMRDFPLVDGHNDLPLVLRQVYQKGLQDVNLRNFSYG
QTSLDRLRDGLVGAQFWSAYVPCQTQDRDALRLTLEQIDLIRRMCAZYSELELVTSAKALND
TQKLACLIGVEGGHSLDNSLSILRTFYMLGVRYLTLHTCNTPWAESSAKGVHSFYNNISGL
TDFGEKVVVAEMNRLGMMVDLSHVSDAVARALEVSQAPVIFSHSAARGVCNSARNVPDDILQ
LLKKNGGVMVMSMGVIQCNPSANVSTVADHFDHIKAVIGSKFIGIGGDYDGAGKFPQGLE
DVSTYPVLIELLSRGWSEEELQGVLRGNLLRVFRQVEKVQEENKWQSPLEDKFPDEQLSSS
CHSDLRSRLRQRQSLTSGQELTEIPIHWTAKLPAKWSVSESSPHDKTHTCPPCPAPELLGGP
SVFLFPPPDKDT
```

bioRxiv preprint doi: <https://doi.org/10.1101/2023.09.04.553312>; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [aCC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).

FIGURE 13

CGCCCAGCGACGTGCGGGCGGCCTGGCCCGGCCCTCCCGGCCGGCTGCGTCCCGGCC
CTGCGCCACCGCCGCCGAGCCGCAGCCCGCCGCCGGCAGCGCCGGCCCC**ATG**CCC
GCCGGCCGCCGGGGCCCCGCCGCCAATCCGCGCGGCCGCCGGCTTGCTGCCCTGCT
GCTGCTGCTCTGCGTCCTCGGGCGCCGCAGGCCGGATCAGGAGCCCACACAGCTGTGATCA
GTCCCCAGGATCCCACGCTTCTCATCGGCTCCTCCCTGCTGGCACCTGCTCAGTGCACGGA
GACCCACCAGGAGCCACCGCCGAGGGCCTCTACTGGACCCCTAACGGCGCCGCTGCC
TGAGCTCTCCCGTGTACTCAACGCCCTCACCTGGCTCTGGCCCTGGCCAACCTCAATGGGT
CCAGGCAGCGGTGGGGACAACCTCGTGTGCCACGCCGTGACGGCAGCATTGGCTGG
TCCTGCCTCTATGTTGGCCTGCCCTGGACGCCAGGGGCCACGGGAGACCTCCTCACACCA
CATGAAGGACTTGACCTGCCCTGGACGCCAGGGGCCACGGGAGACCTCCTCACACCA
ACTACTCCCTCAAGTACAAGCTTAGGTGGTATGCCAGGACAACACATGTGAGGAGTAC
ACAGTGGGGCCCCACTCCTGCCACATCCCCAAGGACCTGGCTCTTTACGCCCTATGAGAT
CTGGGTGGAGGCCACCAACGCCCTGGCTCTGCCCGCTCCGATGTACTCACGCTGGATATCC
TGGATGTGGTGACCACGGACCCCCCGCCGACGTGCACGTGAGCCCGTGGGGGCTGGAG
GACCAGCTGAGCGTGCCTGGGTGTCGCCACCCGCCCTCAAGGATTCTCTTCAAGCAA
ATACCAAGATCCGCTACCGAGTGGAGGACAGTGTGGACTGGAAGGTGGACGATGTGAGCA
ACCAGACCTCCTGCCCTGCCGGCCTGAAACCCGGCACCGTGTACTCGTCAAGTGC
TGCAACCCCTTGGCATCTATGGCTCCAAGAAAGCCGGATCTGGAGTGAGTGGAGGCCACCC
CACAGCCGCCTCCACTCCCCCAGTGAGCGCCGGCCGGCGGGCGTGCACCGC
GGGGCGGAGAGCCGAGCTCGGGCCGGTGCAGCGAGCTCAAGCAGTCCCTGGCTGGCTC
AAGAAGCACCGTACTGCTCCAACCTCAGCTTCCGCCTCTACGACCAAGTGGCGAGCCTGGAT
GCAGAAAGTCGACAAGACCCGCAACCAGGACGAGGGATCCTGCCCTGGCAGACGGGCA
CGCGAGAGGTCTGCCAGA**TAA**GCTGTAGGGCTCAGGCCACCCCTGCCACGTGGAGA
CGCAGAGGCCAACCCAAACTGGGCCACCTCTGATCCACTCAGGGCACCTGAGCCAC
CCTCAGCAGGAGCTGGGTGGCCCTGAGCTCCAACGCCATAACAGCTCTGACTCCCACGT
GAGGCCACCTTGGGTGCACCCAGTGGGTGTGTGTGAGGGTTGGTTGAGTGC
CTAGAACCCCTGCCAGGGCTGGGGTGAGAAGGGAGTCATTACTCCCCATTACCTAGGGCC
CCTCCAAAAGAGTCCTTTAAATAAATGAGCTATTAGGTGCTGTGATTGTGAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAACAAAAAAA

FIGURE 14

><ss.DNA38113

><subunit 1 of 1, 422 aa, 1 stop

><MW: 46302, pI: 9.42, NX(S/T): 6

MPAGRRGPAAQSARRPPPLLPLLLLCLVGAPRAGSGAHTAVISPQDPTLLIGSSLLATCSV
HGDPPGATAEGLYWTNGRRLLPELSRVLNASTLALALANLNGSRQRSGDNLVCHARDGSIL
AGSCLYVGLPPEKPVNISCWSKNMKDLTCRTPGAHGETFLHTNYSLKYLWYGQDNTCEE
YHTVGPHSCHIPKDLALFTPYEIWVEATNRLLGSARSDVLTLDILDVVTTDPPPVDHVSERVGG
LEDQLSVRWVSPPALKDFLFQAKYQIRYRVEDSVDWKVVDDVSNQTSCRLAGLKPGTVYFVQ
VRCNPFGIYGSKKAGIWSEWSHPTAASTPRSERPGPGGACEPRGGEPSGPVRRELKQFLG
WLKKHAYCSNLSFRILYDQWRAMQKSHKTRNQDEGILPSGRRGTARGPAR

Signal sequence.

amino acids 1-30

Transmembrane domain.

amino acids 44-61

N-glycosylation sites.

amino acids 92-96, 104-108, 140-144, 168-172, 292-296, 382-386

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 413-417

N-myristoylation sites.

amino acids 30-36, 37-43, 73-79, 121-127, 179-185, 218-224,
300-306, 317-323, 320-326, 347-353, 355-361, 407-413

Amidation site.

amino acids 3-7, 79-83, 411-415

Growth factor and cytokines receptors family signature 2.

amino acids 325-331

FIGURE 15

FIGURE 16

</usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA34436
<subunit 1 of 1, 300 aa, 1 stop
<MW: 32964, pI: 9.52, NX(S/T): 1
MKFLLDILLLPLLIVCSLESFVKLFIPKRRKSVTGEIVLITGAGHGIGRLTAYEFAKLKSK
LVLWDINKHGLEETAAKCKGLGAKVHTFVVDCSNREDIYSSAKVKAEIGDVSILVNNAGVV
YTSDLFATQDPQIEKTFEVNVLAHFWTTKAFLPAMTKNNHGHIVTVASAAGHVSVPFLAYC
SSKFAAVGFHKTLTDELAALQITGVKTTCLCPNFVNTGFIKNPSTSLGPTLEPEEVVNRLMH
GILTEQKMFIPSSIAFLTTLERILPERFLAVLKRKISVKFDAVIGYKMKAQ

Signal sequence.

amino acids 1-19

Transmembrane domain.

amino acids 170-187

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 30-34, 283-287

N-myristoylation sites.

amino acids 43-49, 72-78, 122-128, 210-216

FIGURE 17

GAATGTTCTCTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAAGGACTG
GGGTGACGGCAGGGCAGGGGGCGCTGGCCGGGAGAAGCGCGGGGGCTGGAGCACCAAA
CTGGAGGGTCCGGAGTAGCGAGGCCCGAAGGAGGCCATGGGGAGCCGGAGGGGGACT
GCGAGAGGACCCGGCGTCCGGCTCCGGTGCAGCGCTATGAGGCCACTCCTCGTCCTGC
TGCTCCTGGCCTGGCGCCGGCTGCCCCCCACTGGACGACAACAAGATCCCCAGCCTCTGC
CCGGGGCACCCGGCCTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCCGGCCG
CGATGGCCCGACGCCCGACGGCGCGACGGCGCCGGGCTCCGGAGAGAAAGGCGAGGGCGGA
GGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGAGGAGAGGCCGGACCCGCGGG
CCCACCGGGCCTGCCGGGAGTGCTCGGTGCCTCCCGCATCCGCCTTCAGGCCAACGCTC
CGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCTTGCCCTCGACCGCGTGCCTGGTAACG
AGCAGGGACATTACGACGCCGTACCGCAAGTTCACCTGCCAGGTGCCTGGGTCTACTAC
TTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTGATCTGGTAAGAATGGCGA
ATCCATTGCCCTTTCTTCAGTTTCGGGGGTGCCAACGCCAGCCTCGCTCTCGGGGG
GGGCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTAC
ATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTCTCCGGATTCTGGTGTACTCCGA
CTGGCACAGCTCCCCAGTCTTGCTTAGTGCCACTGCAAAGTGAGCTATGCTCACTCC
TAGAAGGAGGGTGTGAGGCTGACAACCCAGGTATCCAGGAGGGCTGGCCCCCTGGAATATT
GTGAATGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCTGCTGGCAAGGAATGGGAAC
AGTGGCTGTGCGATCAGGTCTGGCAGCAGGGCAGTGGCTGGATTCTGCCAACGACCA
GAGGAGTGTGCTGTGCTGGCAAGTGTAAAGTCCCCAGTTGCTGGTCCAGGAGGCCACGGT
GGGTGCTCTTCTGGCCTCTGCTCTGGATCCTCCCCACCCCTCCTGCTCCTGG
GCCGGCCCTTTCTCAGAGATCACTCAATAAACCTAAGAACCTCATAAAAAAAAAAAAAA
AAAAAAAAAAAAA

FIGURE 18

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40592
><subunit 1 of 1, 243 aa, 1 stop
><MW: 25298, pI: 6.44, NX(S/T): 0
MRPLLVLLLLGLAAGSPLDDNKIPSLCPGHPGLPGTPGHGSQGLPGRDGRDGRDGAPGAP
GEKGEGRGPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLP
FDRV LVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWP
KPASLSGGAMVRLEPEDQVWVQVGVDYIGIYASIKT DSTFSGFLVYSDWHSSPVFA
```

Signal sequence.

amino acids 1-15

N-myristoylation sites.

amino acids 11-17, 68-74, 216-222

Cell attachment sequence.

amino acids 77-80

FIGURE 19

CTCTTTGTCCACCAGCCCAGCCTGACTCCTGGAGATTGTGAATAGCTCCATCCAGCCTGAG
AAACAAGCGGGTGGCTGAGCCAGGCTGTGCACGGAGCACCTGACGGGCCAACAGACCCAT
GCTGCATCCAGAGACCTCCCCTGGCCGGGGCATCTCCTGGCTGTGCTCCTGGCCCTCCTG
GCACCACTGGCAGAGGTGTGGCCACCCAGCTGCAGGAGCAGGCTCCGATGGCCGGAGCC
CTGAACAGGAAGGAGAGTTCTGCTCCTCCCTGCACAACCGCCTGCGCAGCTGGTCCA
GCCCCCTGCGGCTGACATGCGGAGGCTGGACTGGAGTGACAGCCTGGCCAACTGGCTCAAG
CCAGGGCAGCCCTCTGTGGAATCCAACCCCGAGCCTGGCATCCGGCTGTGGCGCACCTG
CAAGTGGGCTGGAACATGCAGCTGCTGCCGCGGCTTGGCTCCTTGTGAAGTGGTCAG
CCTATGGTTGCAGAGGGCAGCGGTACAGCCACGCCAGGAGAGTGTGCTCGAACGCCA
CCTGCACCCACTACACGCAGCTCGTGTGGCACCTCAAGCCAGCTGGCTGTGGCGGCAC
CTGTGCTCTGCAGGCCAGACAGCGATAAGCCTTGCTGTGCCTACTCCCCGGAGGCAA
CTGGGAGGTCAACGGGAAGACAATCATCCCTATAAGAAGGGTGCCTGGTTCGCTTGCA
CAGCCAGTGTCTCAGGCTGCTCAAAGCCTGGGACCATGCAGGGGGCTCTGTGAGGTCCCC
AGGAATCCTTGTGCATGAGCTGCCAGAACCATGGACGTCTAACATCAGCACCTGCCACTG
CCACTGTCCCCCTGGCTACACGGCAGATACTGCCAAGTGAGGTGCAGCCTGCAGTGTG
ACGGCCGGTTCCGGAGGAGGTGCTCGTGCCTGTGACATCGGCTACGGGGAGGCCAG
TGTGCCACCAAGGTGCATTTCCCTCCACACCTGTGACCTGAGGATCGACGGAGACTGCTT
CATGGTGTCTTCAGAGGCAGACACCTATTACAGAGCCAGGATGAAATGTCAGAGGAAAGGCG
GGGTGCTGGCCCAGATCAAGAGCCAGAAAGTGCAGGACATCCTCGCCTTATCTGGCCGC
CTGGAGACCACCAACGAGGTGACTGACAGTGACTTCGAGACCAGGAACCTCTGGATCGGGCT
CACCTACAAGACGCCAAGGACTCCTCCGCTGGCCACAGGGAGCACCAGGCCTCACCA
GTTTGCCTTGGCAGCCTGACAACCACGGCTGGTGTGGCTGAGTGCTGCCATGGGTTT
GGCAACTGCGTGGAGCTGCAGGCTTCAGCTGCCTCAACTGGAACGACCAGCGCTGCAAAC
CCGAAACCGTTACATCTGCCAGTTGCCAGGAGCACATCTCCGGTGGGCCAGGGTCC
GAGCCTGACCACATGGCTCCCTGCCCTGGAGCACCAGGCTCTGCTTACCTGTCTGC
CCACCTGTCTGGAACAAGGGCCAGGTTAAGACCACATGCCTCATGTCAAAGAGGTCTCAGA
CCTTGCACAATGCCAGAAGTGGCAGAGAGAGGCAGGGAGGCCAGTGAGGGCCAGGGAGTG
AGTGTAGAAGAAGCTGGGCCCTCGCCTGCTTGTGATTGGGAAGATGGCTTCAATTAGA
TGGCGAAGGAGAGGACACGCCAGTGGCCAAAAAGGCTGCTCTTCCACCTGGCCAGAC
CCTGTGGGCAGCGGAGCTCCCTGTGGCATGAACCCACGGGTATTAAATTATGAATCAG
CTGAAAAAAAAAAAAA

FIGURE 20

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44176
<subunit 1 of 1, 455 aa, 1 stop
<MW: 50478, pI: 8.44, NX(S/T): 2

MLHPETSPGRGHLLAVLLALLGTTWAEVWPPQLQEQQAPMAGALNRKESFLLSLHNRLRSWV
QPPAADMRRLDWSDSLAQLAQARAALCGIPTPSLASGLWRTLQVGWNMQLLPAGLASFVEVV
SLWFAEGQRYSHAAGECARNATCTHYTQLVWATSSQLGCRHLCAGQTAIEAFVCAYSPGG
NWEVNGKTIIPYKKGAWSLCTASVSGCFKAWDHAGGLCEVPRNPCRMSQNHGRLNISTCH
CHCPPGYTGRYCQVRCSLQCVHGRFREEECSCVCDIGYGGAQCATKVHFPFHTCDLRIDGDC
FMVSSEADTYYRARMKCQRKGGLAQIKSQKVQDILAFYLGRLETTNEVTDSDFETRNFWIG
LTYKTAKDSFRWATGEHQAFTSFAFGQPDNHGLVWLSAAMFGNCVELQASAAFNWNDQRCK
TRNRYICQFAQEHISRWGP

Signal sequence.
amino acids 1-26

Transmembrane domain.
amino acids 110-124

N-glycosylation sites.
amino acids 144-148, 243-247

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 45-49

N-myristoylation sites.
amino acids 22-28, 99-105, 131-137, 201-207, 213-219, 287-293,
288-294, 331-337, 398-404

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 204-215

EGF-like domain cysteine pattern signature.
amino acids 249-261, 280-292

C-type lectin domain signature.
amino acids 417-442

FIGURE 21

CGGACGCGTGGCTGGCGCTGCAAAGCGTGTCCGCCGGTCCCCGAGCGTCCC CGCC
CGCCCCGCC**ATG**CTCTGCTGCTGGGGCTGTGCCTGGGCTGTCCCTGTGTGGGGTCGCA
GGAAGAGGCCAGAGCTGGGCCACTCTCGGAGCAGGATGGACTCAGGGTCCC GAGGCAAG
TCAGACTGTTGCAGAGGCTGAAAACAAACCTTGATGACAGAATTCTCAGTGAAGTCTACC
ATCATTTCCGTTATGCCTTCACTACGGTTCTGCAGAACATGCTGAACAGAGCTTCTGAAGA
CCAGGACATTGAGTCCAGATGCAGATTCCAGCTGCAGCTTCACTACCAACTTCACTATGC
TTATTGGAGACAAGGTGTATCAGGGCAAATTACAGAGAGAGAAAAGAAGAGTGGTGATAGG
GTAAAAGAGAAAAGGAATAAAACCACAGAAGAAAATGGAGAGAAGGGGACTGAAATATTCA
AGCTTCTGCAGTGATTCCCAGCAAGGACAAAGCCGCCTTTCTGAGTTATGAGGAGCTTC
TGCAGAGGCCCTGGCAAGTACGAGCACAGCATCAGCGTGCAGGGCCAGCAGCTGTCCGG
AGGCTGAGCGTGGACGTGAATATCCTGGAGAGCGCGGGCATCGCATCCCTGGAGGTGCTGCC
GCTTCACAACAGCAGGCAGAGGGCAGTGGCGGGAAAGATGATTCTGGCCTCCCCAT
CTACTGTCATTAACCAAAATGAAACATTGCCAACATAATTTAAACCTACTGTAGTACAA
CAAGCCAGGATTGCCAGAATGGAATTGGAGACTTATCATTAGATATGACGTCAATAG
AGAACAGAGCATTGGGACATCCAGTTCTAAATGGCTATTGTCAGTACTTTGCTCCTA
AAGACCTTCCTCCTTACCCAAGAATGTGGTATTGTCAGTGCCTATGGT
GGAACCAAACCGGCAGACCCAAGGATGCCCTTCAACATTCTCCATGACCTCCGACCCCA
GGACCGTTTCAGTATCATTGGATTTCACCCGATCAAAGTATGGAAGGACCACTTGATAT
CAGTCACTCCAGACAGCATCAGGGATGGGAAAGTGTACATTACCATATGTCACCACGTGGA
GGCACAGACATCAACGGGCCCTGCAGAGGGCAGCAGGCTCTAACAAAGTACGTGGCCA
CAGTGGCATTGGAGACGGAGCGTGTCCCTCATCGTCTCCTGACGGATGGGAAGGCCACGG
TCGGGAGACGCACACCCCTCAAGATCCTCAACAAACACCCGAGAGGCCGCCCAGGCAAGTC
TGCATCTTCACCATTGGCATGGCAACGACGTGGACTTCAGGCTGCTGGAGAAACTGTCGCT
GGAGAACTGTGGCCTCACACGGCGCGTGCACGAGGAGGAGGACGCAGGCTCGCAGCTCATCG
GGTTCTACGATGAAATCAGGACCCCGCTCTCTGACATCCGATCGATTATCCCCCAGC
TCAGTGGTGCAGGCCACCAAGACCCCTGTTCCCAACTACTTCAACGGCTCGGAGATCATCAT
TGCAGGGAAAGCTGGTGGACAGGAAGCTGGATCACCTGACGTGGAGGTACCGCCAGCAACA
GTAAGAAATTCATCATCCTGAAGACAGATGTCCGTGCGGCCCTCAGAAGGCAGGGAAAGAT
GTCACAGGAAGCCCCAGGCCTGGAGGCGATGGAGAGGGGACACCAACCACATCGAGCGTCT
CTGGAGCTACCTCACCACAAAGGAGCTGCTGAGCTCCTGGCTGCAAAGTGA CGATGAACCGG
AGAAGGAGCGGCTGCCAGGGGCCAGGCCCTGGCTGTGAGCTACCGCTTCACTCCC
TTCACCTCCATGAAGCTGAGGGGCCCGTCCCACGCATGGATGGCCTGGAGGAGGCCACGG
CATGTCGGCTGCCATGGGACCCGAACCGGTGGTCAGAGCGTGCAGGAGCTGGCACGCAGC
CAGGACCTTGCTCAAGAAGCCAAACTCCGTAAAAAAAAACAAAACAAAACAAAAAAGA
CATGGGAGAGATGGTGTGTTCTCCACCACTGGGATACGAT**TGAGAAGATGCCACCT**
GCAAGCCAGGAAGACGCCCTCACAGACACCATGTCAGCTGGCACCTGATCTGGACCTC
CCAGCCTCCAGAACTGTGAGAAATAATGTGTTAAGCTAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAA

FIGURE 22

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44192
<subunit 1 of 1, 694 aa, 1 stop
<MW: 77400, pI: 9.54, NX(S/T): 6
MLLLGLCLGLSLCVGSQEEAQSWGHSSSEQDGLRVPRQVRLLQRLKTPLMTEFSVKSTIIS
RYAFTTVSCRMLNRASEDQDIEFQMQIPAAAFITNFTMLIGDKVYQGEITEREKSGDRVKE
KRNKTTEENGEKGTEIFRASAVIPSVDKAFFLSYEELLQRRLGKYEHSISVRPQQQLSGRLS
VDVNILESAGIASLEVLPLHNSRQRGSGRGEDDSGPPPSTVINQNETFANIIFKPTVVQQAR
IAQNGILGDFIIRYDVNREQSIGDIQVLNGYFVHYFAPKDLPLPKNVVFVLDSSASMVGTK
LRQTKDALFTILHDLRPQDRFSIIGFSNRIKVWKDHLISVTPDSIRDGKVIHHMSPTGGTD
INGALQRAIRLLNKYVAHSGIGDRSVSLIVFLTDGKPTVGETHTLKILNNNTREAARGQVCIF
TIGIGNDVDFRLLEKLSLENCGLRRVHEEDAGSQLIGFYDEIRTPLLSDIRIDYPPSSVV
QATKTLFPNYFNGSEIIIAGKLVDRKLDHLHVEVTASNKKFIILKTDVPVRPQKAGKDVTG
SPRPGGDGEGDTNHIERLWSYLTTKELLSSWLQSDDEPEKERLRQRAQALAVSYRFLTPFTS
MKLRGPVPRMDGLEEAHGMSAAMGPEPVVQSVRGAGTQPGPLLKKPNSVKKQNKTKKRHGR
DGVFPLHHLGIR
```

Signal sequence.

amino acids 1-14

N-glycosylation sites.

amino acids 97-101, 127-131, 231-235, 421-425, 508-512, 674-678

Glycosaminoglycan attachment sites.

amino acids 213-217, 391-395

N-myristoylation sites.

amino acids 6-12, 10-16, 212-218, 370-376, 632-638, 638-644

FIGURE 23

CGGACGCGTGGGTGCCGACATGGCGAGTGTAGTGCTGCCGAGCGGATCCCAGTGTGCGGC
GGCAGCGCGGGCGGCGCCTCCGGCTCCGGCTCTGCTGTTGCTCTCCGCCGCGG
CACTGATCCCCACAGGTGATGGCAGAAATCTGTTACGAAAGACGTGACAGTGATCGAGGGA
GAGGTTGCGACCATCAGTTGCCAAGTCATAAGAGTGACGACTCTGTGATTCACTGAA
TCCAACAGGCAGACCATTATTCAGGGACTTCAGGCCTTGAAGGACAGCAGGTTCACTGAA
TGCTGAATTTCTAGCAGTGAACCTCAAAGTATCATTGACAAACGTCTCAATTCTGATGAA
GGAAGATACTTTGCCAGCTCTACCGATCCCCACAGGAAAGTTACACCACCATCACAGT
CCTGGTCCCACCACGTAATCTGATGATCGATATCCAGAAAGACACTGCGGTGGAAGGTGAGG
AGATTGAAGTCAACTGCACTGCTATGCCAGCAAGCCAGCAGACTATCAGGTGGTTCAAA
GGGAACACAGAGCTAAAGCAAATCGGAGGTGGAAGAGTGGTCAGACATGTACACTGTGAC
CAGTCAGCTGATGCTGAAGGTGACAAGGAGGACATGGGTCCCAGTGATCTGCCAGGTGG
AGCACCCCTGCGGTCACTGGAAACCTGCAGACCCAGCGGTATCTAGAAGTACAGTATAAGCCT
CAAGTGCACATTCACTGACTTATCCTCTACAAGGCTTAACCCGGGAAGGGGACGCGCTTGA
GTTAACATGTGAAGCCATCGGGAAAGCCCCAGCCTGTGATGGTAACTTGGGTGAGAGTCGATG
ATGAAATGCCTCAACACGCCGTACTGTCTGGGCCAACCTGTTCAATAACCTAAACAAA
ACAGATAATGGTACATACCGCTGTGAAGCTTCAAACATAGTGGGAAAGCTCACTCGGATTA
TATGCTGTATGTATACGATCCCCCACAACATACCTCCTCCACAACAACCACCA
CCACCACCACCACCACCATCCTTACCATCATCACAGATTCCGAGCAGGTGAAGAAGGC
TCGATCAGGGCAGTGGATCATGCCGTGATCGGTGGCGTGTGGCGGTGGTGGTGGTGGTGGTGG
GCTGTGCTTGCATCATTCTGGGCGCTATTGCGACATAAAGGTACATACTTCAC
ATGAAGCCAAAGGAGCCGATGACGCAGCAGACAGCAGACAGCTATAATCAATGCAGAAGGA
GGACAGAACAACTCCGAAGAAAAGAAAGAGTACTTCATCTTAGATCAGCCTTTGTTCAAT
GAGGTGTCCAACTGGCCCTATTAGATGATAAAGAGACAGTGATATTGG

FIGURE 24

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39518
<subunit 1 of 1, 440 aa, 1 stop
<MW: 48240, pI: 4.93, NX(S/T): 7

MASVVLPSGSQCAAAAAAAPPGLRLLLLLFSAAALIPTGDGQNLFKDVTVIEGEVATISC
QVNKSDDSVIQLLNPNRQTIYFRDFRPLKDSRFQLLNFSSSELKVSLTNVSISDEGRYFCQL
YTDPPQESYTTITVLVPPRNLMIDIQKDTAVEGEEIEVNCTAMASKPATTIRWFKGNTTELKG
KSEVEEWSDMYTVTSQLMLKVHKEDDGVPVICQVEHPAVTGNLQTQRYLEVQYKPQVHIQMT
YPLQGLTREGDALELTCEAIGKPQPVMVTWVRVDDEMPQHAVLSGPNLFINNLNKTDNGTYR
CEASNIVGKAHSDYMLYVYDPPTTIPPTTTTTTTTILTIITDSRAGEEGSIRAVDH
AVIGGVVAVVVFAMLCLLIILGRYFARHKGTYFTHEAKGADDAADADTAIINAEGGQNNSEE
KKEYFI

Signal sequence.

amino acids 1-36

Transmembrane domain.

amino acids 372-393

N-glycosylation sites.

amino acids 65-69, 99-103, 111-115, 163-167, 302-306, 306-310,
430-434

Tyrosine kinase phosphorylation sites.

amino acids 233-240, 319-328

N-myristoylation sites.

amino acids 9-15, 227-233, 307-313, 365-371, 376-382, 402-408,
411-417, 427-433, 428-432

FIGURE 25

GGGGCGGGTGGACGCGGACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCGA
CCGCCAGAAAAGACTGAGGCCGCGCCTGCCCGCCCGCTCCCTGCGCCGCCGCCCTC
CCGGGACAGAAAGATGTGCTCCAGGGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCT
GGGCCTGGGTGCAGGGCTGCCATCCGGCTGCCAGTGCAAGCCAGCACAGACAGTCTTCT
GCACTGCCGCCAGGGACACGGTCCCCGAGACGTGCCACCCGACACGGTGGGGCTGTAC
GTCTTGAGAACGGCATCACCATGCTCGACGAAGCAGCTTGCCGGCTGCCGGCCTGCA
GCTCTGGACCTGTACAGAACAGATGCCAGCCTGCCCTGCCCGCCCTGCTGCTGCTGG
ACCTCAGCCACAACAGCCTCTGCCCTGGAGCCGGCATCCTGGACACTGCCAACGTGGAG
GCGCTGCCGCTGGCTGGCTGGGCTGCAGCAGCTGGACGAGGGCTTCAAGCCGTTGCC
AACCTCCACGACCTGGATGTGTCGACAACCAGCTGGAGCGAGTGCCACCTGTGATCCGAG
GCCTCCGGGCCTGACGCCCTGCCCTGGCCGGCAACACCCGATTGCCAGCTGCCGGCCC
GAGGACCTGCCGCCCTGGCTGCCCTGCAGGAGCTGGATGTGAGCAACCTAACGCTGCC
CCTGCCTGGCGACCTCTGCCCTCTCCCCGCCCTGCCCTGCGCTGCTGGCAGCTGCCCAACC
CCTCAACTGCGTGTGCCCTGAGCTGGTTGGCCCTGGTGCAGGAGAGGCCACGTACA
CTGGCCAGCCCTGAGGAGACCGCCTGCCACTTCCCGCCAACGACTGGCCGGCTGCTCCT
GGAGCTTGAACCGCCGACTTGGCTGCCAGCCACCACAGCCACAGTGCCACCCA
CGAGGCCCGTGGTGCAGGAGCCACAGCCTGTCTTAGCTTGGCTCCTACCTGGCTTAGC
CCCACAGGCCGCCACTGAGGCCAGGCCCTCACGCCACCGACTGTAGGGCC
TGTCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCCAATGGGGCACATGCCACCTGG
GGACACGGCACCACCTGGCGTCTTGCCCCGAAGGCTTACGGGCTGTACTGTGAGAGC
CAGATGGGGCAGGGACACGGCCAGCCCTACACCAGTCAGGCCAGGCCACCACGGTCCCT
GACCCCTGGCATCGAGCCGGTGAGCCCCACCTCCCTGCCGTGGGCTGCAGCGTACCTCC
AGGGAGCTCGTGCAGCTCAGGAGCCTCCGTCACCTATGCAACCTATCGGGCCCTGAT
AAGCGGCTGGTACGCTGCGACTGCCCTGCTCGCTGAGTACACGGTACCCAGCTGCC
GCCAACGCCACTTACTCCGTCTGTGTCATGCCCTTGGGGCCGGGCGGGTGCAGGGCG
AGGAGGCCTGCCGGGAGGCCATACACCCCCAGCCGTCACTCCAACCACGCCAGTCACC
CAGGCCCGAGGGCAACCTGCCCTCCATTGCCGCCCTGGCGCGGTGCTCCTGGC
CGCGCTGGCTGCCGGTGGGGCAGCCTACTGTGTCGGGGGGCGGGCATGGCAGCAGCGG
CTCAGGACAAAGGGCAGGTGGGGCAGGGCTGGGCCCTGGAACACTGGAGGGAGTGAAGGTC
CCCTGGAGCCAGGCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGTCTGA
GTGTGAGGTGCCACTCATGGCTTCCAGGGCTGCCAGTCACCCCTCACGCCAAAGC
CCTACATCTAAGGCAGAGAGACAGGGCAGCTGGGCCGGCTCTCAGCCAGTGAGATGGC
CAGCCCCCTCTGCCACACCACGTAAGTTCTCAGTCCCAACCTGGGGATGTGTCAGA
CAGGGCTGTGACCAACAGCTGGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAG
ATGCTGTGCCAGCTGACGAGCCCTAACGTCCCCAGAACCGAGTGCCCTATGAGGACAGTGT
CCGCCCTGCCCTCCGCAACGTGCAGTCCCTGGCACGGCGGGCCCTGCCATGTGCTGGTAAC
GCATGCCCTGGGCCCTGCTGGCTCTCCACTCAGGCCAGGACCTGGGGCCAGTGAAGGAAG
CTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGCGGCTGTGACTCTAGTCTGGCCCTTAA
AAGCGAAGGAACAAAGAAACTGGAAAGGAAGATGCTTAGGAACATGTTGCTTTAA
AATATATATATTTATAAGAGATCCCTTCCATTATCTGGGAAGATGTTTCAAACCTC
AGAGACAAGGACTTGGTTTGTAAGACAAACGATGATGAAGGCCTTGTAAAGAAAAA
ATAAAAAAAAAA

FIGURE 26

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44804

<subunit 1 of 1, 598 aa, 1 stop

<MW: 63030, pI: 7.24, NX(S/T): 3

MCSRVPLLLPLLLLALGPGVQGCPSCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFEN
GITMLDASSFAGLPGQLLDSLQNQIASLRLPRLLLLDLSHNSLLALEPGILDANVEALRL
AGLGLQQLDEGLFSRRLRNLDLVDSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDLA
GLAALQELDVSNLSQLALPGDLSGLFPRLLLAAARNPFNCVCPLSWFGPWVRESHVTLASP
EETRCHFPPKNAGRLLLELDYADFGCPATTTATVPTTRPVVREPTALSSSLAPTWLSPTAP
ATEAPSPPPSTAPPTVGVPQPQDCPPSTCLNGGTCHLGTRHHLACLCPEGFTGLYCESQMGQ
GTRPSPTPVTPRPPRSLTGLIEPVSPSLRVGLQRYLQGSSVQLRSLRLTYRNLSGPDKRLV
TLRLPASLAEYTWTQLRPNATYSVCVMPGLPGRVPEGEACGEAHTPPAVHSNHAPVTQARE
GNLPLLIAPALAAVLLAALAAGAACVRRGRAMAAAQDKGQVPGAGPLELEGVKVPLEP
GPKATEGGEALPSGSECEVPLMGFPGPGLQSPLHAKPYI

Signal sequence.

amino acids 1-23

Transmembrane domain.

amino acids 501-522

N-glycosylation sites.

amino acids 198-202, 425-429, 453-457

Tyrosine kinase phosphorylation site.

amino acids 262-270

N-myristoylation sites.

amino acids 23-29, 27-33, 112-118, 273-279, 519-525, 565-571

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

EGF-like domain cysteine pattern signature.

amino acids 355-367

Leucine zipper pattern.

amino acids 122-144, 194-216

FIGURE 27

GGCACTAGGACAACCTTCTCCCTCTGCACCCTGCCGTACCCCTACCGCCCCGCCACC
TCCTTGCTACCCCCTCTGAAACCACAGCTGTTGGCAGGGTCCCCAGCTC**ATGCCAGCCTC**
ATCTCCTTCTTGCTAGCCCCAAAGGGCCTCCAGGCAACATGGGGGCCAGTCAGAGAGC
CGGCACTCTCAGTTGCCCTCTGGTTGAGTTGGGGGGCAGCTCTGGGGCCGTGGCTTGCC
ATGGCTCTGCTGACCCAAACAAACAGAGCTGCAGAGCCTCAGGAGAGAGGTGAGCCGGCTGCA
GGGGACAGGAGGCCCCCTCCCAGAATGGGAAGGGTATCCCTGGCAGAGTCTCCGGAGCAGA
GTTCCGATGCCCTGGAAGCCTGGAGAAATGGGGAGAGATCCCGAAAAGGAGAGCAGTGCTC
ACCCAAAAACAGAAGCAGCACTCTGCTGCACCTGGTCCCATTAAACGCCACCTCAA
GGATGACTCCGATGTGACAGAGGTGATGTGGCAACCAGCTCTTAGCGTGGAGAGGCCTAC
AGGCCAAGGATATGGTGTCCGAATCCAGGATGCTGGAGTTATCTGCTGTATAGCCAGGTC
CTGTTCAAGACGTGACTTCAACCATGGTCAGGTGGTCTCGAGAAGGCCAAGGAAGGCA
GGAGACTCTATTCCGATGTATAAGAAGTATGCCCTCCACCCGGACCGGGCTACAACAGCT
GCTATAGCGCAGGTGTCTCCATTACACCAAGGGATATTCTGAGTGTCTAATTCCCCGG
GCAAGGGCGAAACTAACCTCTCCACATGGAACCTCCTGGGTTGTGAAACTGT**GATT**
GTGTTATAAAAAGTGGCTCCAGCTTGGAAAGACCAGGGTGGTACATACTGGAGACAGCCAA
GAGCTGAGTATATAAGGAGAGGAATGTGCAGGAACAGAGGCATCTCCTGGGTTGGCTC
CCCGTTCCCTCACTTTCCCTTCATTCCCACCCCTAGACTTGATTTACGGATATCTG
CTTCTGTTCCCCATGGAGCTCCG

FIGURE 28

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52722
<subunit 1 of 1, 250 aa, 1 stop
<MW: 27433, pI: 9.85, NX(S/T): 2
MPASSPFLLAPKGPPGNMGGPVREPALSVALWLSWGAALGAVACAMALLTQQTELQLRREV
SRLQGTGGPSQNGEYPWQSLPEQSSDALEAWENGERSRKRAVLTQKQKKQHSVLHLVPIN
ATSKDDSDVTEVMWQPALRRGRGLQAQGYGVRIQDAGVYLLYSQVLFQDVTFMGQVVSREG
QGRQETLFRCIRSMPSHPDRAYNSCYSAGVFHLHQGDILSVIIPRARAKLNLSPHGTFLGFVKL
```

Signal sequence.

amino acids 1-40

N-glycosylation site.

amino acids 124-128

Tyrosine kinase phosphorylation site.

amino acids 156-164

N-myristoylation site.

amino acids 36-42, 40-46, 179-185, 242-248

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 34-45

FIGURE 29

CACTTTCTCCCTCTCTCCTTACTTTGAGAAACCGCGCTTCCGCTTCTGGTCGCAGAGAC
CTCGGAGACCGCGCCGGGAGACGGAGGTGCTGTGGGTGGGGGACCTGTGGCTGCTCGTA
CCGCCCCCCCACCTCCTCTGCACTGCCGTCTCCGAAGACCTTCCCCTGCTCTGTT
TCCTTCACCGAGTCTGTGCATGCCCGGACCTGGCGGGAGGAGGCTTGCCGGCGGGAGA
TGCTCTAGGGCGCGCGGGAGGAGCAGGCCGGGGACGGAGGGCCGGCAGGAAG**ATGGC**
TCCCGTGGACAGGGACTCTTGTCTGGCGTACTGCCTGCTCCTGCTTGCCTCTGGCCTGGT
CCTGAGTCGTGTGCCCATGTCCAGGGGAACAGCAGGAGTGGAGGGACTGAGGAGCTGC
CGTCGCCTCCGGACCATTGCCGAGAGGGCTGAAGAACAAACATGAAAAATACAGGCCAGTCAG
GACCAGGGCTCCCTGCTTCCCGTGCTGCGCTGCTGTGACCCCGTACCTCATGTACCC
GGCGACCGCCGTGCCCATGAGATCAACATCACTATCTGAAAGGGAGAAGGGTGACCGCGGAG
ATCGAGGCCTCCAAGGGAAATATGGCAAAACAGGCTCAGCAGGGCCAGGGCCACACTGGA
CCCCAAAGGGCAGAAGGGCTCCATGGGGGCCCTGGGAGCGGTGCAAGGCCACTACGCCGC
CTTTTCGGTGGGCCGGAAGAAGCCCATGCAACAGCAACCCTACTACAGACGGTGATCTCG
ACACGGAGTTCTGAACCTCTACGACCACTTCAACATGTTCACCGGCAAGTTCTACTGCTAC
GTGCCCGGCTCTACTTCTCAGCCTCAACGTGACACCTGGAACCAGAAGGAGACCTACCT
GCACATCATGAAGAACGAGGAGGAGGTGGTATCTTGTTCGCGCAGGTGGCGACCGCAGCA
TCATGCAAAGCCAGAGCCTGATGCTGGAGCTGCGAGAGCAGGACAGGTGTGGTACGCCTC
TACAAGGGCGAACGTGAGAACGCCATCTCAGCGAGGAGCTGGACACCTACATCACCTCAG
TGGCTACCTGGTCAAGCACGCCACCGAGCC**TAG**CTGGCGGCCACCTCCTCTCGCC
ACCTTCCACCCCTGCGCTGTGCTGACCCCACCGCCTCTCCCGATCCCTGGACTCCGACTC
CCTGGCTTGGCATTCACTGAGACGCCCTGCAACACAGAAAAGCCAAGCGATCGGTGCTCC
CAGATCCCAGCCTCTGGAGAGAGCTGACGGCAGATGAAATCACCAGGGGGGACCCGC
GAGAACCCCTGGACCTTCCCGGCCCTCTGCAACACATCCTCAAGTGACCCGCACGGC
GAGACGCGGGTGGCGCAGGGCGTCCCAGGGTGGCACCGCGGCTCCAGTCCTGGAAATA
ATTAGGCAAATTCTAAAGGTCTCAAAAGGAGCAAAGTAAACCGTGGAGGACAAAGAAAAGGG
TTGTTATTTGTCTTCCAGCCAGCCTGCTGGCTCCAAAGAGAGAGGCCTTTCAAGTTGAG
ACTCTGCTTAAGAGAAGATCCAAGTTAAAGCTCTGGGTGAGGGAGGGGGCGGGGAGG
AAACTACCTCTGGCTTAATTCTTTAAGCCACGTAGGAACCTTCTGAGGGATAGGTGGACC
CTGACATCCCTGTGGCCTGCCAAGGGCTCTGCTGGTCTTCTGAGTCACAGCTGCGAGGT
GATGGGGCTGGGCCAGGGCTAGCCTGCCAGAGGGACAGCTGAGCCCCCTGCCTTGGC
TCCAGGTTGTTAGAACGAGCCGAAGGGCTCCTGACAGTGGCCAGGGACCCCTGGTCCCCA
GGCCTGCAGATGTTCTATGAGGGCAGAGCTCCTGGTACATCCATGTGCGCTCTGCTCC
ACCCCTGTGCCACCCAGAGCCCTGGGGGTGGTCTCCATGCCTGCCACCTGGCATCGGCT
TTCTGTGCCCTCCACACAAATCAGCCCCAGAAGGCCGGGCTTGGCTTGTGTTTT
TATAAAACACCTCAAGCAGCACTGCAGTCTCCATCTCCTCGTGGCTAAGCATCACCCT
CCACGTGTGTTGTTGGCAGCAAGGCTGATCCAGACCCCTCTGCCCTGGACTGCCCT
CATCCAGGCCTCTGACCACTGAGAGGGCTTTCTAGGCTTCAGAGCAGGGAGAG
CTGGAAGGGGCTAGAAAGCTCCGCTTGTCTGTTCTCAGGCTCTGAGCCTCAGTCCTG
AGACCAGAGTCAAGAGGAAGTACACGTCCAATCACCGTGTAGGATTCACTCTCAGGAGC
TGGGTGGCAGGAGAGGCAATAGCCCCGTGGCAATTGCAGGACAGCTGGAGCAGGGTTGCG
GTGCTCCACGGTGTCTGCCCTGCCATGGCCACCCAGACTCTGATCTCCAGGAACCC
ATAGCCCCCTCTCCACCCATGTTGATGCCAGGGTCACTCTGCTACCCGCTGGGCC
CCCAAACCCCGCTGCCCTCTTCCCTCCCCCATCCCCCACCTGGTTTGACTAAAGAACTT
TTCCTCTCTGGGCCCTGGCTGCCGGATCTGGGTCCCTAAGTCCCTCTCTTTAAAGAACTT
CTGCGGGTCAGACTCTGAAGCCGAGTTGCTGTGGCGTGCCGGAGCAGAGGCCACACTC
GCTGCTTAAGCTCCCCAGCTTTCCAGAAAACATTAAACTCAGAATTGTGTTTCAA

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41234

><subunit 1 of 1, 281 aa, 1 stop

><MW: 31743, pI: 6.83, NX(S/T): 1

MGSRGQGLLLAYCLLLAFASGLVLSRVPHVQGEQQEWEGTEELPSPPDHAERAEEQHEKYRP
SQDQGLPASRCLRCCDPGTSMYPATAVPQINITILKGEKGDRGDRGLQGKYGKTGSAGARGH
TGPKGQKGSMGAPGERCKSHYAAFSVGRKKPMHSNHYYQTVIFDTEFVNLYDHFNMF TGKFY
CYVPGLYFFSLNVHTWNQKETYLHIMKNEEVVILFAQVGDRSIMQSQSLMELREQDQVWV
RLYKGERENAIFSEELDTYITFSGYLVKHATEP

Signal sequence.

amino acids 1-25

N-glycosylation site.

amino acids 93-97

N-myristoylation sites.

amino acids 7-13, 21-27, 67-73, 117-123, 129-135

Amidation site.

amino acids 150-154

Cell attachment sequence.

amino acids 104-107

FIGURE 31

GGGGAGGCATCCGCTCGGGTCCTCGCCGAGACCCCCCGCGCGATTGCCGGTCTTCCCGGG
GCGCAGACAGAGCTGTCTCGCACCTGGATGGCAGCAGGGCGCGGGGCTCTCGACGCCA
GAGAGAAATCTCATCATCTGTGCAGCCTCTAAAGCAAACATAAGACCAAGAGGGAGGATTAT
CCTTGACCTTGAAGACAAAACACTAAACTGAAATTAAAATGTTCTCGGGGGAGAAGGGAG
CTTGACTTACACTTGGTAATAATTGCTCCTGACACTAAGGCTGCTGCTAGTCAGAATT
GCCTCAAAAGAGTCTAGAAGATGTTGTCATTGACATCCAGTCATCTCTTCTAAGGGAATC
AGAGGCAATGAGCCGTATATACTTCAACTCAAGAAGACTGCATTAATTCTTGCTGTTCAAC
AAAAAACATATCAGGGACAAAGCATGTAATTGATGATCTCGACACTCGAAAAACAGCTA
GACAACCCAAC TGCTACCTATTCTGTCCTAACGAGGAAGCCTGTCATTGAAACCAGCA
AAAGGACTTATGAGTTACAGGATAATTACAGATTTCATCTTGACCAGAAATTGCCAAG
CCAAGAGTTACCCAGGAAGATTCTCTTACATGCCATTTCACAAGCAGTCACCCCC
TAGCCCACATCATCACACAGATTATTCAAAGCCCACCGATATCTCATGGAGAGACACACTTCT
CAGAAGTTGGATCCTCAGATCACCTGGAGAAACTATTAAAGATGGATGAAGCAAGTGCCA
GCTCCTTGCTTATAAGGAAAAGGCCATTCTCAGAGTTCACAAATTCTGATCAAGAAA
TAGCTCATCTGCTGCCTGAAAATGTGAGTGCGCTCCAGCTACGGTGGCAGTTGCTTCTCCA
CATACCACCTCGGCTACTCCAAAGCCCACCCCTTACCCACCAATGCTTCAGTGACACC
TTCTGGGACTTCCCAGCCACAGCTGCCACCACAGCTCCACCTGTAACCACGTCACTTCTC
AGCCTCCCACGACCTCATTCTACAGTTTACACGGCTGCGGCTACACTCCAAGCAATG
GCTACACAGCAGTCTGACTACCACCTTCAGGCACCTACGGACTCGAAAGCAGCTAGA
AACCATACCGTTACAGAAATCTCAAACCTTAACATTGAACACAGGAAATGTGTATAACCC
CTGCACTTCTATGTCAAATGTGGAGTCTCCACTATGAATAAAACTGCTTCTGGAGG
AGGGAGGCCAGTCCAGGCAGTFCTCTCCAGGGCAGTGTCCAGAAAATCAGTACGGCCTTCC
ATTGAAAATGGCTTCTATGGGTCCTGCTCTGGTGTCTGTTCTGGTGTAGGCC
TCGTCCTCCTGGGTAGAATCCTTCGGAATCACTCCGCAGGAAACGTTACTCAAGACTGGAT
TATTGATCAATGGGACTATGTGGACATTAAGGATGGAACTCGGTGTCTCTTAATTCTT
TAGTAACCAGAAGCCAAATGCAATGAGTTCTGCTGACTGCTAGTCTTAGCAGGAGGTTG
TATTGAAAGACAGGAAAATGCCCTCTGCTTCTGCTTCTTCTTGGAGACAGAGTCTT
GCTCTGTTGCCAGGCTGGAGTGCAGTAGCACGATCTCGGCTCTACCGCAACCTCCGCTC
CTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCTAAGTATCTGGATTACAGGCATGTGCCA
CCACACCTGGGTGATTTGTATTTTAGTAGAGACGGGTTTACCATGTTGGTCAGGCTG
GTCTCAAACCTCTGACCTAGTGTACCCACCCCTCTGGCCTCCAAAGTGTGGATTACAGG
CATGAGCCACCACAGCTGGCCCCCTCTGTTATGTTGGTTTGAGAAGGAATGAAGTG
GGAACCAAATTAGGTAATTGGTAATCTGTCTAAACATTAGCTAAAACAAAGCTCT
ATGTAAAGTAATAAAAGTATAATTGCCATATAAAATTCAAACCTGGCTTTATGCAA
GAAACAGGTTAGGACATCTAGGTTCCAATTCAATTGATCTGCTTCTTATGATTCCAGA
TGTGTTATGATCTAATTGATGTTGCTTTCTTATGATTGATCTTAAAACATTATT
CCAGATGTAGTTCCATTCAAATTAAATATTGAATAATCTTGTACTCAA

FIGURE 32

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45410
><subunit 1 of 1, 431 aa, 1 stop
><MW: 46810, pI: 6.45, NX(S/T): 6
MFFGGEGSLTYTLVIICFLTLRLSASQNCLKKSLEDVVIDIQSSLSKGIRGNEPVYTSTQED
CINSCCSTKNISGDKACNLMIFDTRKTARQPNCYLFFCPNEEACPLKPAKGLMSYRIITDFP
SLTRNLPSQELPQEDSLLHGQFSQAVTPLAHHTDYSKPTDISWRDTLSQKFGSSDHLEKLF
KMDEASAQLLAYKEKGHSQSSQFSSDQEIAHLLPENVSALPATVAVASPHTTSATPKPATLL
PTNASVTPSGTSQPQLATTAPPVTTVTSQPPTTLISTVFTRAATLQAMATTAVLTTFQAP
TDSKGSLETIPFTEISNLTLNTGNVYNPTALSMSNVESSTMNKASWEGREASPGSSSQGSV
PENQYGLPFEKWLLIGSLLFGVLFLVIGLVLLGRILSESLRRKRYSRLDYLINGIYVDI
```

Signal sequence.

amino acids 1-25

Transmembrane domain.

amino acids 384-405

N-glycosylation sites.

amino acids 72-76, 222-226, 251-255, 327-331, 352-356

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 415-419

Tyrosine kinase phosphorylation site.

amino acids 50-57

N-myristoylation sites.

amino acids 4-10, 48-54, 315-321

FIGURE 33

GC GG CAC CTG G AAG **ATG** CGCC CATT GG CT GG TG GC CT GCT CA AG GT GG T GTT CG TGG T CTT C
GC CT CTT GT GT GC CT GG T ATT CGGG GT AC CT GCT CG CAG AG CT CATT CC AG AT GC ACCC CT
GT CC AG TG CT GC CT AT AG C AT CC CG CAG C AT CGGG AG AG GG CT GT CC CT CAA AG CT CC AG TCC
CC AAA AGG CAAA AT GT GACC ACT GG ACT CC CT GCCC AT CT GAC AC CT AT GC CT AC AG GT TA
CT CAG CGG AGG TGG CAG AAG CA AGT AC G CCA AA AT CT GCT T GAGG ATA AC CT ACT T AT GGG
AG AAC AG CT GGG AA AT GTT GCC AG AG GA AT AA AC AT TG CC ATT GT CA ACT AT GT AACT GGG A
AT GT GAC AG CA AC AC GAT GT TT GAT AT GT AT GA AGG CG AT A ACT CT GG ACC GAT GAC AA AG
TT T ATT CAG AGT GCT GCT CAAA AT CC CT GCT CT TC AT GG T GAC CT AT GAC GAC GGA AGC AC
AAG ACT GA AT A AC GAT G CCA AGA AT G CC AT AGA AGC ACT TT GG AAGT AA AGA AAT CAG G A A C A
TG AA ATT CAG GT CT AG CT GGG T ATT ATT G CAG CAA AAG G CT T GG AACT CC CT CCG A A ATT
CAG AGA GAA AAG AT CA ACC ACT CT GAT GCT AAG A AC A CAG AT ATT CT GG CT GG C CT G CAG A
GAT CC AG AT AGA AGG CT G C AT ACC CAA AAG A AC G A AG C **TGA** C ACT G CAG GG C CT GAG TAA AT
GT GT T CT GT AT AA ACA A AT G CAG CT GG A AT CG CT CA AG A AT CT T ATT TT CT AA AT CC A A C A
GCC C AT ATT GAT GAG T ATT TGG TT GT T G TAA ACC A AT G A AC AT TT G C T A G T GT T AT CA
AAT CTT GG T AC G CAG T ATT TT AT ACC AGT ATT T AT GT A G T G A A G AT GT C A AT TAG CAG G A
AACT AAA AT G A AT GG A A ATT CTT AAAA AAAAAA A

FIGURE 34

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46777
><subunit 1 of 1, 235 aa, 1 stop
><MW: 25982, pI: 9.09, NX(S/T): 2
MRPLAGLLKVVFVVFASLCAWYSGYLLAELIPDAPLSSAAYSIRSIGERPVLKAPVPKRQK
CDHWTPCPSDTYAYRLLSGGGRSKYAKICFEDNLLMGEQLGNVARGINIAIVNYVTGNVTAT
RCFDMDYEGDNGPMTKFIQSAAPKSLLFMVTYDDGSTRLNNDAKNAIEALGSKEIRNMKFRS
SWVFIAAKGLELPSEIQREKINHSDAKNNRYSGWPABEQIEGCIPKERS
```

Signal sequence.

amino acids 1-20

N-glycosylation sites.

amino acids 120-124, 208-212

Glycosaminoglycan attachment site.

amino acids 80-84

N-myristoylation sites.

amino acids 81-87, 108-114, 119-125